

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Maruyama, Ichiro  
Maruyama, Hiroko  
Brenner, Sydney
- (ii) TITLE OF INVENTION: LAMBDROID BACTERIOPHAGE VECTORS FOR  
EXPRESSION OF FOREIGN PROTEINS
- (iii) NUMBER OF SEQUENCES: 15
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: The Scripps Research Institute, Office of  
Patent Counsel
  - (B) STREET: 10666 North Torrey Pines Road, TPC8
  - (C) CITY: La Jolla
  - (D) STATE: CA
  - (E) COUNTRY: USA
  - (F) ZIP: 92037
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US
  - (B) FILING DATE: 05-AUG-1994
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Fitting, Thomas
  - (B) REGISTRATION NUMBER: 34,163
  - (C) REFERENCE/DOCKET NUMBER: TSRI432.0
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 619-554-2937
  - (B) TELEFAX: 619-554-6312

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 246 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Pro Val Pro Asn Pro Thr Met Pro Val Lys Gly Ala Gly Thr Thr  
1 5 10 15  
Leu Trp Val Tyr Lys Gly Ser Gly Asp Pro Tyr Ala Asn Pro Leu Ser  
20 25 30  
Asp Val Asp Trp Ser Arg Leu Ala Lys Val Lys Asp Leu Thr Pro Gly  
35 40 45  
Glu Leu Thr Ala Glu Ser Tyr Asp Asp Ser Tyr Leu Asp Asp Glu Asp  
50 55 60  
Ala Asp Trp Thr Ala Thr Gly Gln Gly Gln Lys Ser Ala Gly Asp Thr  
65 70 75 80  
Ser Phe Thr Leu Ala Trp Met Pro Gly Glu Gln Gly Gln Gln Ala Leu  
85 90 95  
Leu Ala Trp Phe Asn Glu Gly Asp Thr Arg Ala Tyr Lys Ile Arg Phe  
100 105 110  
Pro Asn Gly Thr Val Asp Val Phe Arg Gly Trp Val Ser Ser Ile Gly  
115 120 125  
Lys Ala Val Thr Ala Lys Glu Val Ile Thr Arg Thr Val Lys Val Thr  
130 135 140  
Asn Val Gly Arg Pro Ser Met Ala Glu Asp Arg Ser Thr Val Thr Ala  
145 150 155 160  
Ala Thr Gly Met Thr Val Thr Pro Ala Ser Thr Ser Val Val Lys Gly  
165 170 175  
Gln Ser Thr Thr Leu Thr Val Ala Phe Gln Pro Glu Gly Val Thr Asp  
180 185 190  
Lys Ser Phe Arg Ala Val Ser Ala Asp Lys Thr Lys Ala Thr Val Ser  
195 200 205  
Val Ser Gly Met Thr Ile Thr Val Asn Gly Val Ala Ala Gly Lys Val  
210 215 220  
Asn Ile Pro Val Val Ser Gly Asn Gly Glu Phe Ala Ala Val Ala Glu  
225 230 235 240  
Ile Thr Val Thr Ala Ser

245

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGTGTGGAGC TCTACCCTTT C

21

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCCTGTAAT AAGCGGCCGC AGCT

24

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGGCCGCTT ATTACAGG

18

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 910 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16..897

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAATGTGAGG ACGCT ATG CCT GTA CCA AAT CCT ACA ATG CCG GTG AAA GGT	51
Met Pro Val Pro Asn Pro Thr Met Pro Val Lys Gly	
1 5 10	
GCC GGG ACC ACC CTG TGG GTT TAT AAG GGG AGC GGT GAC CCT TAC GCG	99
Ala Gly Thr Thr Leu Trp Val Tyr Lys Gly Ser Gly Asp Pro Tyr Ala	
15 20 25	
AAT CCG CTT TCA GAC GTT GAC TGG TCG CGT CTG GCA AAA GTT AAA GAC	147
Asn Pro Leu Ser Asp Val Asp Trp Ser Arg Leu Ala Lys Val Lys Asp	
30 35 40	
CTG ACG CCC GGC GAA CTG ACC GCT GAG TCC TAT GAC GAC AGC TAT CTC	195
Leu Thr Pro Gly Glu Leu Thr Ala Glu Ser Tyr Asp Asp Ser Tyr Leu	
45 50 55 60	
GAT GAT GAA GAT GCA GAC TGG ACT GCG ACC GGG CAG GGG CAG AAA TCT	243
Asp Asp Glu Asp Ala Asp Trp Thr Ala Thr Gly Gln Gly Gln Lys Ser	
65 70 75	
GCC GGA GAT ACC AGC TTC ACG CTG GCG TGG ATG CCC GGA GAG CAG GGG	291
Ala Gly Asp Thr Ser Phe Thr Leu Ala Trp Met Pro Gly Glu Gln Gly	
80 85 90	
CAG CAG GCG CTG CTG GCG TGG TTT AAT GAA GGC GAT ACC CGT GCC TAT	339
Gln Gln Ala Leu Leu Ala Trp Phe Asn Glu Gly Asp Thr Arg Ala Tyr	
95 100 105	
AAA ATC CGC TTC CCG AAC GGC ACG GTC GAT GTG TTC CGT GGC TGG GTC	387
Lys Ile Arg Phe Pro Asn Gly Thr Val Asp Val Phe Arg Gly Trp Val	
110 115 120	
AGC AGT ATC GGT AAG GCG GTG ACG GCG AAG GAA GTG ATC ACC CGC ACC	435
Ser Ser Ile Gly Lys Ala Val Thr Ala Lys Glu Val Ile Thr Arg Thr	
125 130 135 140	

GTG AAA GTC ACC AAT GTG GGA CGT CCG TCG ATG GCA GAA GAT CGC AGC Val Lys Val Thr Asn Val Gly Arg Pro Ser Met Ala Glu Asp Arg Ser 145 150 155	483
ACG GTA ACA GCG GCA ACC GGC ATG ACC GTG ACG CCT GCC AGC ACC TCG Thr Val Thr Ala Ala Thr Gly Met Thr Val Thr Pro Ala Ser Thr Ser 160 165 170	531
GTG GTG AAA GGG TAG AGC TGG CCT GTT AGG CCC ACT CCG ACC CCG ACC Val Val Lys Gly * Ser Trp Pro Val Arg Pro Thr Pro Thr Pro Thr 175 180 185	579
ACT CCC ACC CCG ACT CCC ACC CCG ACC CCG ACC CCG ACT CCG ACC GTT Thr Pro Thr Pro Thr Pro Thr Pro Thr Pro Thr Pro Thr Pro Thr Val 190 195 200	627
GGG CCA ATT GTC ACA CAG GAA ACA GCT ATG ACC ATG ATT ACG CCA AGC Gly Pro Ile Val Thr Gln Glu Thr Ala Met Thr Met Ile Thr Pro Ser 205 210 215 220	675
TTG CAT GCC TGC AGG TCG ACT CTA GAG GAT CCC CGG GTA CCG AGC TCG Leu His Ala Cys Arg Ser Thr Leu Glu Asp Pro Arg Val Pro Ser Ser 225 230 235	723
AAT TCA CTG GCC GTC GTT TTA CAA CGT CGT GAC TGG GAA AAC CCT GGC Asn Ser Leu Ala Val Val Leu Gln Arg Arg Asp Trp Glu Asn Pro Gly 240 245 250	771
GTT ACC CAA CTT AAT CGC CTT GCA GCA CAT CCC CCT TTC GCC AGC TGG Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro Pro Phe Ala Ser Trp 255 260 265	819
CGT AAT AGC GAA GAG GCC CGC ACC GAT CGC CCT TCC CAA CAG TTG CGC Arg Asn Ser Glu Glu Ala Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg 270 275 280	867
AGC CTG AAT GGC GAA TGG CGC CTG TAA TAAGCGGCCG CAGCTC Ser Leu Asn Gly Glu Trp Arg Leu * 285 290	910

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 293 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:



275

280

285

Glu Trp Arg Leu \*  
290

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTCTGCAGCA CAAGCTCAAC CTTA

24

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAGAATTCTT TACATACTGG AATAAGAG

28

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Val Val Lys Gly

1

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTGGTGAAG GGTAGAGCTC CACACTG

27

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CAGTGTGGAG CTCTACCCTT TCACCAC

27

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Ser Ser Leu Asp Pro Gly Pro Ser Thr Asn Ser  
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

00852020.050697  
009050.02025830



(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGGTAGAGCT CAAGCTTGGA TCCGGGCCCCG TCGACGAATT C

41

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ser Trp Pro Val Gly Pro Ile Val Thr Gln Glu Thr Ala Met Thr Met  
1 5 10 15

Ile Thr Pro Ser Leu His Ala Cys Arg Ser Thr Leu Glu Asp Pro Arg  
20 25 30

Val Pro Ser Ser Asn Ser  
35

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

GGGTAGAGCT GGCCTGTTGG GCCAATTGTC ACACAGGAAA CAGCTATGAC CATGATTACG 60  
CCAAGCTTGC ATGCCTGCAG GTCGACTCTA GAGGATCCCC GGGTACCGAG CTCGAATTC 119